二項分布の平均・分散・標準偏差

確率変数 X が二項分布 $B(n,\ p)$に従うとき，平均 $E(X)$，分散 $V(X)$，標準偏差 $\sigma(X)$は

$\boxed{1}$ $\boldsymbol{E(X)=np}$　　$\boxed{2}$ $\boldsymbol{V(X)=npq}$　　$\boxed{3}$ $\boldsymbol{\sigma(X)=\sqrt{npq}}$

ただし　$q=1-p$

証明 $\boxed{1}$

X の確率分布は次の表のようになる。

X	0	1	2	……	r	……	n	計
確率 P	${}_n\mathrm{C}_0\, q^n$	${}_n\mathrm{C}_1\, pq^{n-1}$	${}_n\mathrm{C}_2\, p^2q^{n-2}$	……	${}_n\mathrm{C}_r\, p^rq^{n-r}$	……	${}_n\mathrm{C}_n\, p^n$	1

これより，X の平均 $E(X)$は

$$E(X)=0\times {}_n\mathrm{C}_0\, q^n+1\times {}_n\mathrm{C}_1\, pq^{n-1}+2\times {}_n\mathrm{C}_2\, p^2q^{n-2}+\cdots\cdots+r\times {}_n\mathrm{C}_r\, p^rq^{n-r}+\cdots\cdots+n\times {}_n\mathrm{C}_n\, p^n$$
$$=\sum_{k=0}^{n} k\cdot {}_n\mathrm{C}_k\, p^kq^{n-k}$$

ここで　${}_n\mathrm{C}_r=\dfrac{n!}{r!\,(n-r)!}$　であり，二項定理により

$$(q+p)^n={}_n\mathrm{C}_0\, q^n+{}_n\mathrm{C}_1\, q^{n-1}p+\cdots\cdots+{}_n\mathrm{C}_r\, q^{n-r}p^r+\cdots\cdots+{}_n\mathrm{C}_n\, p^n$$
$$=\sum_{k=0}^{n} {}_n\mathrm{C}_k\, p^kq^{n-k}=\sum_{k=0}^{n}\frac{n!}{k!\,(n-k)!}p^kq^{n-k}$$

であり，$(q+p)^{n-1}=\displaystyle\sum_{k=0}^{n-1}\frac{(n-1)!}{k!\,(n-k-1)!}p^kq^{n-k-1}$　$\cdots\cdots(*)$　である。

$E(X)=\displaystyle\sum_{k=0}^{n} k\cdot {}_n\mathrm{C}_k\, p^kq^{n-k}$ について，$k=0$ のとき $k\cdot {}_n\mathrm{C}_k\, p^kq^{n-k}=0$ であるから

$$E(X)=\sum_{k=0}^{n} k\cdot {}_n\mathrm{C}_k\, p^kq^{n-k}=\sum_{k=1}^{n} k\cdot {}_n\mathrm{C}_k\, p^kq^{n-k}$$

これを変形すると

$$E(X)=\sum_{k=1}^{n} k\cdot {}_n\mathrm{C}_k\, p^kq^{n-k}=\sum_{k=1}^{n} k\cdot\frac{n!}{k!\,(n-k)!}p^kq^{n-k}=\sum_{k=1}^{n}\frac{n!}{(k-1)!\,(n-k)!}p^kq^{n-k}$$

ここで，$k-1=l$ とおくと，$k=n$ のとき $l=n-1$ であるから

$$E(X)=\sum_{l=0}^{n-1}\frac{n!}{l!\,(n-l-1)!}p^{l+1}q^{n-l-1}=\sum_{l=0}^{n-1}\frac{n\cdot(n-1)!}{l!\,(n-l-1)!}p\cdot p^lq^{n-l-1}=np\sum_{l=0}^{n-1}\frac{(n-1)!}{l!\,(n-l-1)!}p^lq^{n-l-1}$$

$(*)$から　$E(X)=np(q+p)^{n-1}$　　$q=1-p$ より　$E(X)=np\cdot 1^{n-1}=np$

証明 2

公式 $V(X)=E(X^2)-\{E(X)\}^2$ を用いる。$E(X)=np$ であるから，$E(X^2)$を求めればよい。

$$E(X^2)=\sum_{k=0}^{n} k^2\cdot {}_n\mathrm{C}_k\, p^k q^{n-k}=\sum_{k=1}^{n} k^2\cdot \frac{n!}{k!\,(n-k)!}p^k q^{n-k}$$

$k=0$ のとき
$k^2\cdot {}_n\mathrm{C}_k\, p^k q^{n-k}=0$ であるから。

ここで，$k!=k\cdot(k-1)\cdot(k-2)!$ であるから，
$k^2=k(k-1)+k$ と考える。

$$\begin{aligned}
E(X^2)&=\sum_{k=1}^{n}\{k(k-1)+k\}\cdot\frac{n!}{k!\,(n-k)!}p^k q^{n-k}\\
&=\sum_{k=1}^{n}k(k-1)\cdot\frac{n!}{k!\,(n-k)!}p^k q^{n-k}+\sum_{k=1}^{n}k\cdot\frac{n!}{k!\,(n-k)!}p^k q^{n-k}\\
&=\sum_{k=2}^{n}\frac{n!}{(k-2)!\,(n-k)!}p^k q^{n-k}+np\\
&=n(n-1)p^2\sum_{k=2}^{n}\frac{(n-2)!}{(k-2)!\,(n-k)!}p^{k-2} q^{n-k}+np
\end{aligned}$$

$k=1$ のとき
$$k(k-1)\cdot\frac{n!}{k!\,(n-k)!}p^k q^{n-k}=0,$$
$$\sum_{k=1}^{n}k\cdot\frac{n!}{k!\,(n-k)!}p^k q^{n-k}=np$$
であるから。

ここで，$k-2=m$ とおくと，$k=n$ のとき $m=n-2$ であるから

$$E(X^2)=n(n-1)p^2\sum_{m=0}^{n-2}\frac{(n-2)!}{m!\,(n-m-2)!}p^m q^{n-m-2}+np$$

二項定理により，$\displaystyle\sum_{m=0}^{n-2}\frac{(n-2)!}{m!\,(n-m-2)!}p^m q^{n-m-2}=(q+p)^{n-2}=1$ であるから

$$E(X^2)=n(n-1)p^2+np$$

よって　$V(X)=E(X^2)-\{E(X)\}^2=n(n-1)p^2+np-(np)^2=n^2p^2-np^2+np-n^2p^2=np(1-p)=npq$

証明 3

$\sigma(X)=\sqrt{V(X)}$ より　$\sigma(X)=\sqrt{npq}$

ポイント　平均$E(X)$の計算式を，二項定理$\displaystyle\sum_{k=0}^{n-1}\frac{(n-1)!}{k!\,(n-k-1)!}p^k q^{n-k-1}=(q+p)^{n-1}$ が代入できるように変形する。

分散 $V(X)=E(X^2)-\{E(X)\}^2$ では，$k^2=k(k-1)+k$ と考えると二項定理の公式が適用できる。